

IDENTIFIKASI AKSARA BALI KE HURUF LATIN DENGAN MENGGUNAKAN KLASIFIKASI TEMPLATE MATCHING DAN K-NEAREST NEIGHBOR (K-NN)

Kadek Ayu Silvia Yuliaratih¹, Joko Haryatno², Koredianto Usman³

¹Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

Abstrak

Aksara Bali berasal dari India yang dibawa oleh para pelajar dari India ke Indonesia. Aksara ini mengalami banyak penyesuaian hingga menjadi huruf tradisional masyarakat Bali. Aksara Bali berpangka l pada huruf Pallawa dan aksara Dewanagari. Aksara ini mirip dengan aksara Jawa. Perbedaananya hanya pada bentuk lekukan bentuk huruf. Aksara Bali telah mengalami masa kejayaannya sebagai bahasa dan aksara resmi di kerajaan-kerajaan yang ada di Bali. Banyak karya tulis, terutama dalam bentuk lontar, yang diciptakan. Lontar-lontar ini sampai kini masih merupakan sumber pengetahuan pokok dari ilmu-ilmu pengobatan tradisional, menentukan saat yang baik untuk melakukan berbagai kegiatan.

Seiring perkembangan waktu, aksara Bali mulai kurang diminati oleh generasi muda. Hal ini disebabkan karena kesulitan untuk mengenali bentuk atau pola dari suku kata dasar aksara tersebut. Suku kata dasar aksara Bali hampir memiliki struktur yang sama, sehingga di dalam pembacaannya akan menemukan kesulitan mengidentifikasi suku kata dasar.

Pada tugas akhir ini dikembangkan suatu aplikasi untuk menkonversikan aksara Bali pada suatu data gambar yang berisi karakter dan berasal dari pemindahan hardcopy (scanning). Proses ekstraksi ciri menggunakan template matching, dimana template matching ini digunakan untuk menjelaskan bagaimana mengenali kembali bentuk atau huruf. Sedangkan klasifikasi cirri menggunakan metode K-Nearest Neighbor.

Pengujian dilakukan dengan mengambil sampel-sampel dari suatu data yang berasal dari scanning. Hasil akurasi terbaik terjadi pada ekstraksi ciri. Secara keseluruhan akurasi sistem yang dibuat sebesar 75% dengan waktu yang berbeda-beda pada setiap karakter.

Kata Kunci : aksara Bali, Template matching, K-nearest neighbor, Scanning.

Telkom
University

Abstract

Balinese ancient alphabet was originally from India and spread by Indian students during their studies in Indonesia. Up until now, this set of characters had gone through several modification phases in order to receive a strong acceptance to become Balinese traditional characters. Balinese ancient alphabet refers to Pallawa as well as Dewanagari characters. It similarly looks like Javanese alphabet but the major difference is only at its unique shape/curve. Balinese ancient alphabet had reached its glorious summit as a written formal letter and language among people in Bali during kingdom age. There were a lot of written documents as well as ancient scripts created using this set of characters and surprisingly, these scripts are still currently being used as a fundamental reference for traditional medicine and also to determine appropriate days for performing certain activities.

Nowadays, Balinese ancient alphabet is less interested by the youth since it is difficult to recognise its basic words' shapes. Balinese ancient alphabet's basic word has similar structure between one and another. Therefore, there will be a tricky problem encountered in identifying the basic word(s).

This Final Year Project develops an application in order to convert Balinese ancient alphabet in an image data that contains characters taken out from hardcopied-materials through scanning processes. The process of unique character extract implements template matching where it is used to explain the process of re-recognising the shape or character. In addition, unique character classification process relies on K-Nearest Neighbor method.

Testing is performed by picking up numerous samples from a scanned data and the best accuracy is shown at the unique character extract process. Overall, system accuracy is set at 75% with the time different for each character.

Keywords : Balinese ancient alphabet, Template Matching, K-Nearest Neighbor, Scanning.



BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Bahasa daerah merupakan kekayaan budaya yang patut dilestarikan. Salah satunya adalah aksara Bali. Aksara atau huruf Bali biasanya digunakan untuk menuliskan bahasa Bali. Seiring perkembangan waktu, aksara Bali mulai kurang diminati oleh generasi muda. Hal ini disebabkan karena kesulitan untuk mengenali bentuk atau pola dari suku kata dasar aksara tersebut. Suku kata dasar aksara Bali hampir memiliki struktur yang sama, sehingga di dalam pembacaannya akan menemukan kesulitan mengidentifikasi suku kata dasar.

Pada tugas akhir ini dikembangkan suatu aplikasi untuk menkonversikan aksara Bali menggunakan *template matching* dan *k-nearest neighbor*, dimana *template matching* nantinya digunakan untuk menjelaskan bagaimana mengenali kembali bentuk atau huruf. Sedangkan klasifikasi ciri menggunakan metode *k-nearest neighbor*. Untuk melakukan analisis dan konversi aksara Bali akan menggunakan Pengolahan Citra Digital. Pengolahan Citra Digital merupakan proses yang menerima masukan berupa citra, melakukan pengolahan terhadapnya, dan menghasilkan keluaran yang berbentuk citra juga. Hal penting yang perlu dikerjakan menyangkut aksara Bali yaitu aturan-aturan yang terdapat dalam proses penulisan huruf atau aksara Bali itu sendiri.

Aplikasi yang dibangun ini berupa aksara Bali yang akan diterjemahkan ke dalam aksara latin. Diharapkan metode yang digunakan dapat dijadikan referensi sebagai salah satu metode pengidentifikasian karakter yang handal, dan aplikasi yang dihasilkan nantinya memiliki tingkat akurasi lebih dari 75 persen.

1.2 Tujuan

Dari permasalahan yang diteliti, maka tujuan dari penulisan tugas akhir ini adalah untuk :

1. Merancang dan membuat program aplikasi yang berfungsi untuk mengenali karakter aksara Bali melalui pengolahan citra dengan menggunakan *k-nearest neighbor*.

2. Menganalisa permormansi program yang dibuat dengan parameter tingkat akurasi sehingga output yang dihasilkan memuaskan.

1.3 Perumusan Masalah

Hal-hal yang akan di analisa dalam Tugas Akhir ini adalah:

1. Segmentasi yang digunakan untuk pemecahan karakter
2. Proses ekstraksi ciri menggunakan *template matching* dan *k-nearest neighbor*.
3. Proses pembelajaran terhadap input-input sampel karakter.
4. Akurasi / ketepatan sistem dalam mengenali karakter aksara Bali.
5. Kecepatan system mulai dari input gambar sampai output akhir yang berupa teks pembacaan.

1.4 Batasan Masalah

Aplikasi yang dibangun ini akan memiliki *fitur* dan *interface* yang menarik. Dalam pembuatan aplikasi ini penulis memberikan batasan masalah yang diantaranya :

1. Citra yang akan diolah adalah aksara Bali dengan format file berupa JPEG
2. Aplikasi ini menggunakan klasifikasi *template matching* dan *k-nearest neighbor algorithm* yang nantinya akan digunakan untuk melakukan klasifikasi suku kata dasar.
3. Implementasi menggunakan bahasa pemrograman MATLAB 7.8 (R2009A).

1.5 Hasil yang diharapkan

Diharapkan penelitian ini menghasilkan aplikasi identifikasi aksara Bali yang dapat mengenali karakter pada suatu file gambar yang berisi karakter dengan tingkat akurasi yang tinggi.

1.6 Metodologi Penelitian

Metode penelitian Tugas Akhir ini meliputi beberapa tahapan, antara lain sebagai berikut :

1. Studi literatur
mengumpulkan data dan mempelajari konsep dasar dan teori-teori yang digunakan untuk sistem konversi aksara bali ke huruf latin menggunakan klasifikasi *template matching* dan *k-nearest neighbor*.
2. Studi pengembangan aplikasi
Bertujuan untuk menentukan metodologi pengembangan sistem yang digunakan dengan pendekatan terstruktur dan melakukan analisis perancangan.
3. Implementasi program aplikasi
Merancang program simulasi yang akan digunakan untuk aplikasi sistem konversi aksara bali ke aksara latin pada citra digital dan mengimplementasikan dengan menggunakan lingkungan pemrograman matlab 7.

1.7 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam pembuatan tugas akhir ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini menjelaskan mengenai Latar Belakang Masalah, Identifikasi Masalah, Batasan Masalah, Maksud dan Tujuan, Metodologi Penelitian dan Sistematika Penulisan.

BAB II LANDASAN TEORI

Bab ini membahas konsep dasar mengenai citra digital, pengolahan citra digital, filter median, operasi morfologi

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini membahas bagaimana proses perancangan program simulasi guna mengkonversikan aksara bali ke aksara latin.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini menguraikan hasil dan analisa dari program simulasi yang telah dibuat.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan hasil dan saran dari Tugas Akhir ini.



BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil analisis terhadap pengujian yang dilakukan pada sistem konversi aksara Bali dengan *Template matching* dan *K-nearest neighbor (K-NN)*, maka dapat ditarik kesimpulan sebagai berikut :

1. Kondisi citra dan hasil yang didapatkan dari hasil *cropping* berpengaruh terhadap pola ciri yang dihasilkan.
2. *Threshold* sangat berpengaruh dalam keberhasilan identifikasi aksara. Jika nilai *threshold* semakin kecil citra yang dihasilkan semakin tipis dan mulai pecah, sedangkan jika nilai semakin besar, citra akan semakin tebal dan memungkinkan citra satu dengan yang lainnya berdempet.
3. nilai *variabel k* (jumlah tetangga terdekat yang digunakan sebagai acuan saat melakukan *voting* dalam penentuan kelas aksara) berpengaruh terhadap tingkat akurasi klasifikasi citra dengan metode *K-Nearest Neighbor*
4. Tingkat akurasi klasifikasi berfluktuasi tergantung jumlah *neighbor* terdekat (nilai *k*) yang digunakan dan tidak ada kecenderungan turun atau naik seiring berubahnya nilai *k*.
5. Pemilihan data latih yang dimasukkan dalam proses pelatihan mempengaruhi tingkat akurasi pada sistem.
6. Klasifikasi pengidentifikasian aksara ke huruf latin membutuhkan waktu yang berbeda-beda dalam setiap pengidentifikasiannya tergantung pada banyaknya karakter dan tanda baca yang terkait di dalamnya

5.2 Saran

1. Untuk kedepannya diharapkan penggunaan proses segmentasi citra untuk bisa memisahkan huruf yang terlalu dekat menggunakan deteksi tepi.
2. Pengembangan aksara Bali dapat mencakup kelas yang lebih banyak lagi.
3. Pengambilan data dilakukan dengan kondisi yang sama seperti menggunakan *scanner*.
4. Menggunakan ukuran data latih yang lebih baik yang dapat menghasilkan ciri yang berbeda untuk masing-masing kelas.
5. Pada proses klasifikasi dapat menggunakan metode yang lain sehingga bisa dibandingkan hasilnya



DAFTAR PUSTAKA

- [1]. Maher A. Sid Ahmed. 1995. "*Image Processing: Theory, Algorithm and Architecture*". McGrawHill.
- [2]. Abdia Away, Gunaidi. 2006. "*The Shortcut of MATLAB Programming*".
Bandung : Penerbit Informatika Bandung.
- [3]. Klein, Dan. (2008). "*Artificial Intelligence – kernels and clustering*".
UC Berkeley.
- [4]. Indyk, Piotr. "*Algorithms for Nearest neighbor search*". MIT.
- [5]. Shahabi, c., tang, L-a., xing, s. "*indexing land surface for efficient kNN query*". University of southern California.
- [6]. Chen, yong-sheng. (2001). "*Fast algorithm for nearest neighbor search based on a lower bound tree*". NTNU CSIE.
- [7] Kenneth R. Castleman. 1996. "*Digital Image Processing*". Prentice Hall
- [8] Teknomo, Kardi. *K-Nearest Neighbor Tutorial*. 2006
<http://people.revoledu.com/kardi/tutorial/KNN/>
- [9] Serra, J. 1982. *Image Analysis and Mathematical Morphology*.
Academic Press, INC. London
- [10] Young, Gerbrands, Vilet. 1998. *Fundamental of Image Processing*.
Delft University of Technology
- [11] Baldock, Graham. 2000. *Image Processing and Analysis - A Practical Approach*. Oxford University Press. New York
- [12] Jahne, Bernd. 2002. *Digital Image Processing 5th Ed*. Springer.